

# Genome-scale analysis for new metabolic engineering procedures

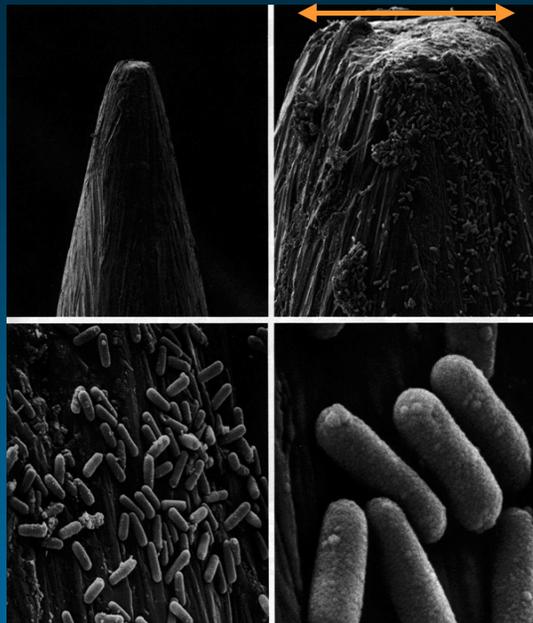
NSF/ME 2003

Bernhard Palsson  
Department of Bioengineering  
UCSD

## Outline

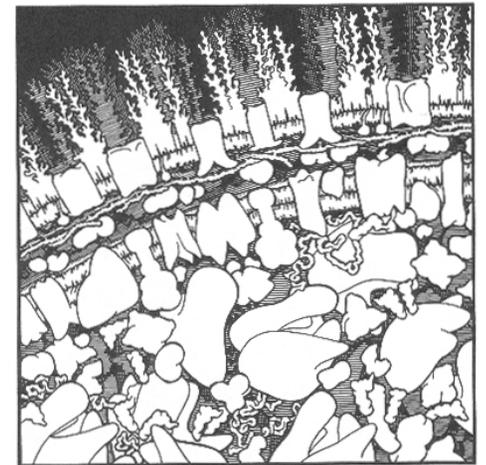
- ¥ Progress with constraints-based models
- ¥ Predicting adaptive evolution
- ¥ Model-driven systems biology

*E. coli*  
on a pin



Alberts et al, 1st edition

The intracellular environment is very complex and counter-intuitive



From The Machinery of Life,  
David S. Goodsell,  
Springer-Verlag, New York, 1993.

Viscosity  $\sim 1000 \times \text{H}_2\text{O}$   
Pressure (osmotic)  $\sim 500 \text{ atm}$   
Electrical gradient  $\sim 300,000 \text{ V/cm}$

# Modern Modeling Approaches

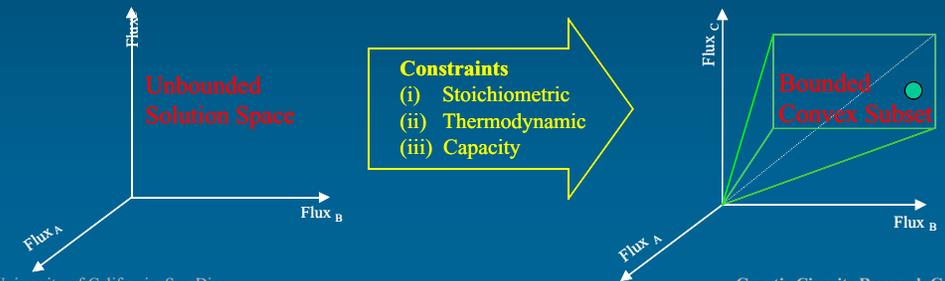
- ⚡ Need to integrate diverse data types (genomic, transcriptomic, proteomic, metabolomic, phenomic,...)
- ⚡ Must be easily scalable to cell or genome-scale
- ⚡ Account for inherent biological uncertainty

# Constraints-based Analysis



*How often have I said to you that when you have eliminated the impossible, whatever remains, however improbable, must be the truth?*

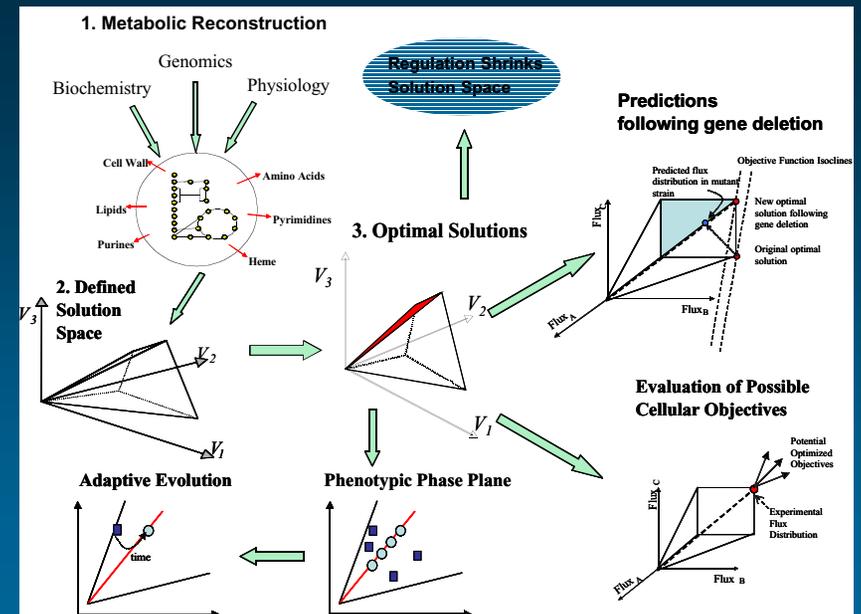
—Sherlock Holmes, A Study in Scarlet



## I. Update on Constraints-Based Modeling

- ⚡ Method development
- ⚡ Reconstruction of regulatory networks
- ⚡ Integration of multiple data types

## Process of Constraints-based Model Building and Analysis

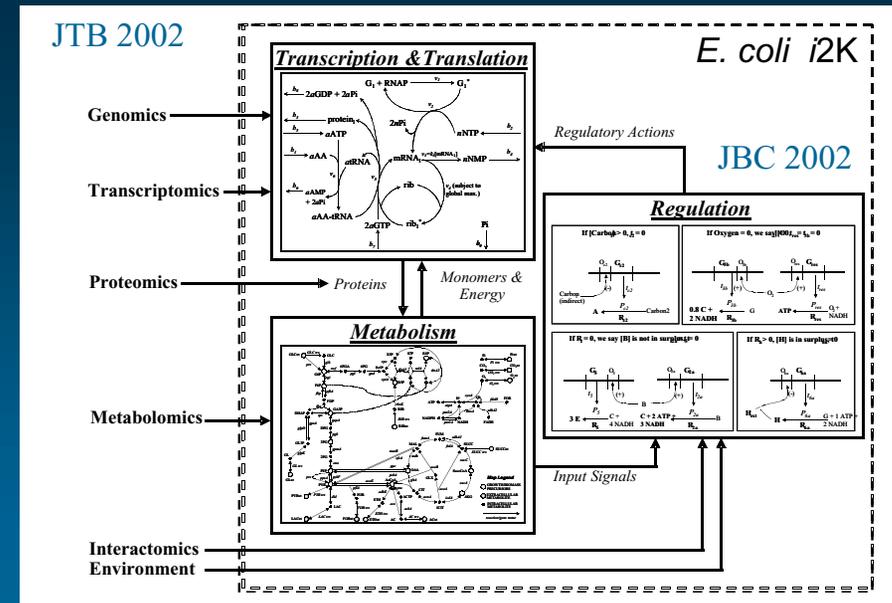


# Constraints-Based Models

| Organism                            | Year | Genes | Reactions | Metabolites |
|-------------------------------------|------|-------|-----------|-------------|
| <i>Escherichia coli (core)</i>      | 1990 | 24    | 14        | 17          |
| <i>Escherichia coli</i>             | 1993 | 250   | 146       | 118         |
| <i>Escherichia coli</i>             | 1998 | 306   | 317       | 305         |
| <i>Haemophilus influenzae</i>       | 1999 | 362   | 488       | 343         |
| <i>Escherichia coli (v1.0)</i>      | 2000 | 695   | 720       | 436         |
| <i>Helicobacter pylori</i>          | 2002 | 291   | 388       | 339         |
| <i>Escherichia coli (regulated)</i> | 2002 | 149   | 113       | 63          |
| <i>Saccharomyces cerevisiae</i>     | 2002 | 957   | 1294      | 801         |
| <i>Escherichia coli (v1.1)</i>      | 2003 | 906   | 1043      | 594         |

Other metabolic networks have been reconstructed for: *Bacillus subtilis*, *Streptococcus pneumoniae*, *Pseudomonas aeruginosa*, *Geobacter sulfurreducens*, *Mycobacterium tuberculosis*, *Anabaena*, and *Plasmodium falciparum*.

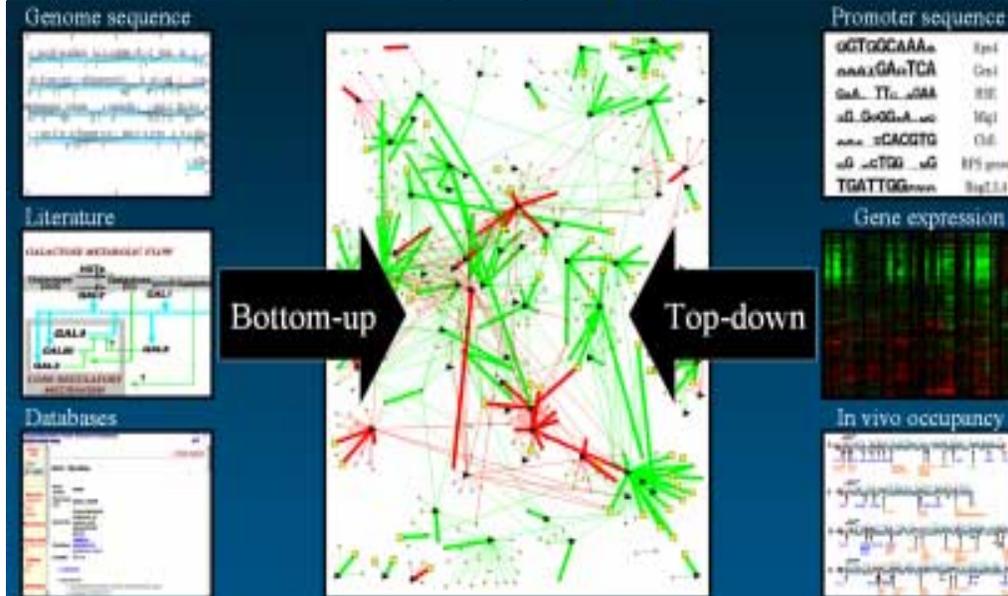
# Integrated whole cell models



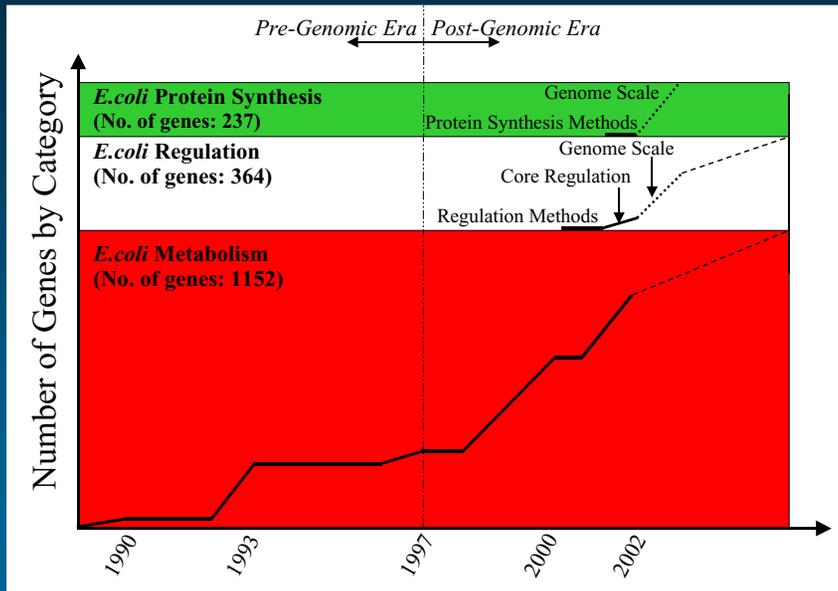
# Reconstruction of Regulatory Networks: Reconciling Bottom-Up and Top-Down Data

| Network Parameters      | <i>E. coli</i> | <i>S. cerevisiae</i> |
|-------------------------|----------------|----------------------|
| Regulatory Genes        | 123            | 109                  |
| Target Genes            | 762            | 418                  |
| Regulatory Interactions | 1468           | 945                  |

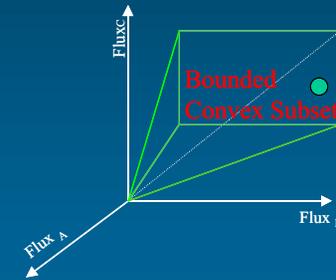
# Reconstructing regulatory networks



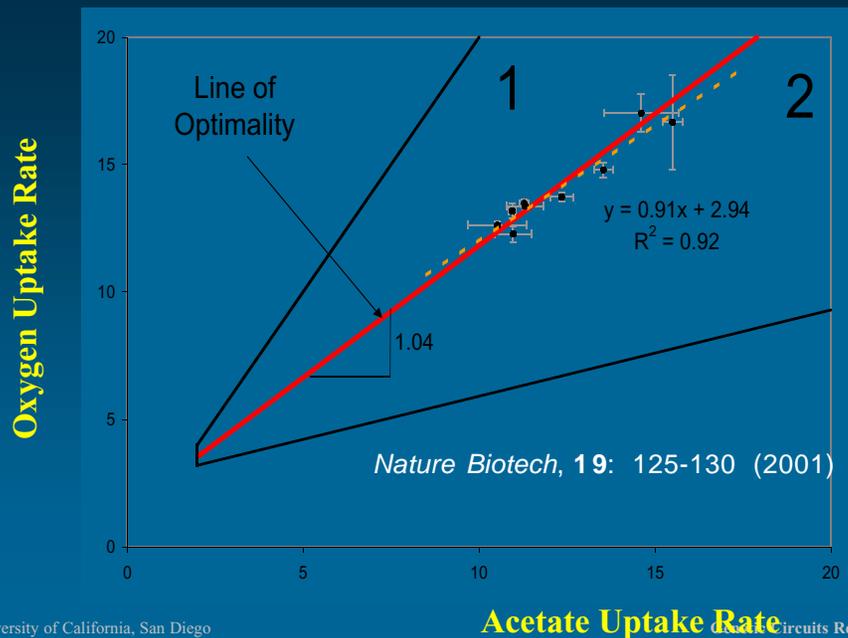
# History of Constraints-Based Models of *E. coli*



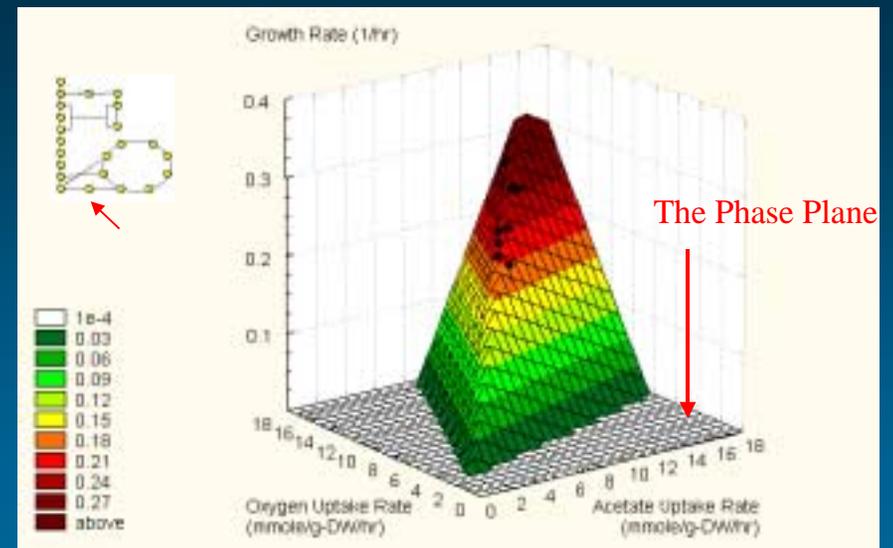
# II. Adaptive Evolution of *E. coli* and *in silico* Predictions



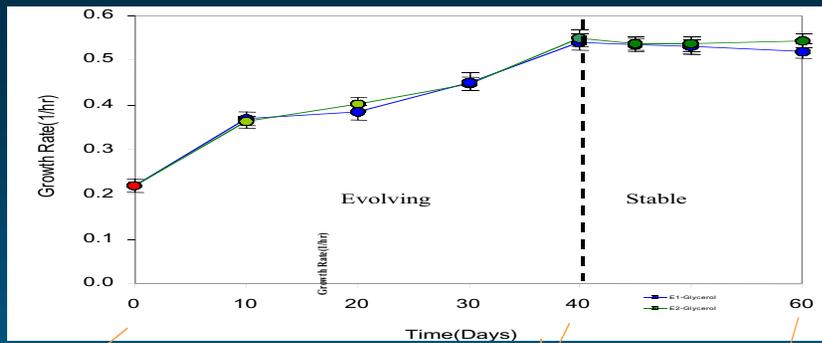
# Growth on acetate: *in silico* vs. *in vivo* results



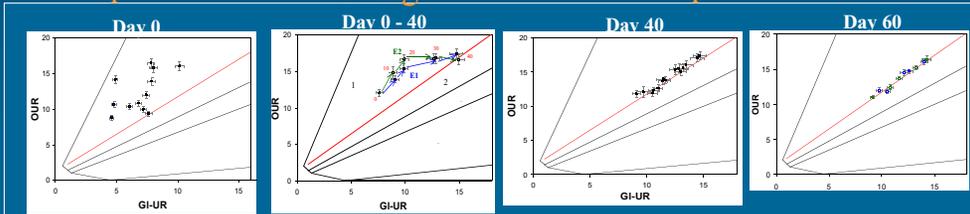
# Growth on Acetate: Life on the edge



# E. coli can Optimize its Growth Rate on Glycerol



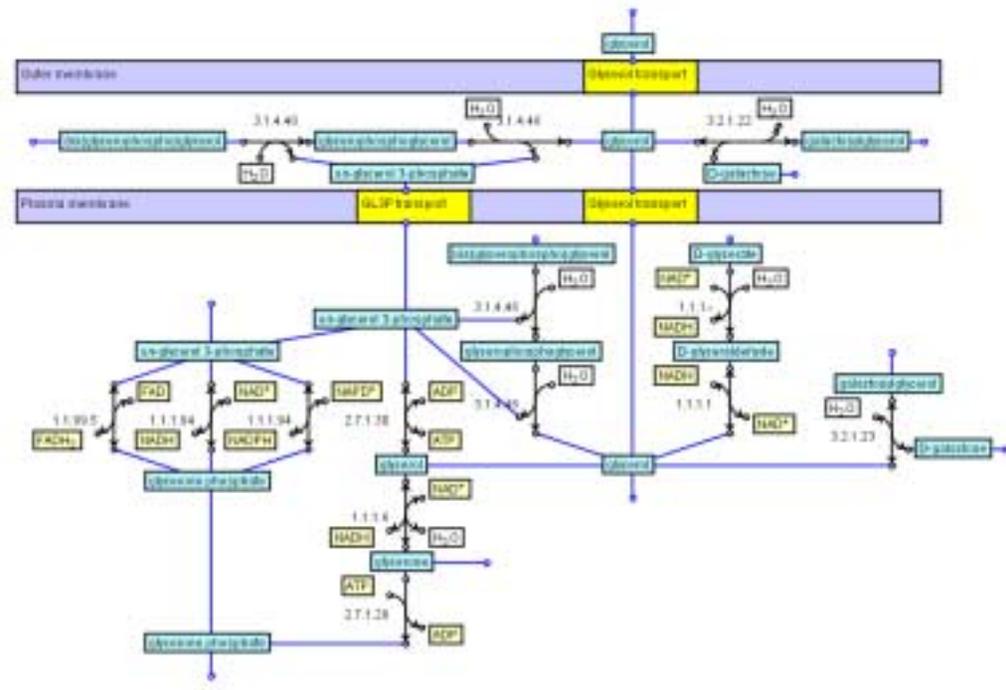
pre during post



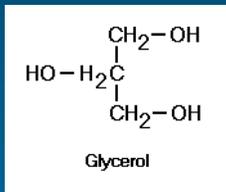
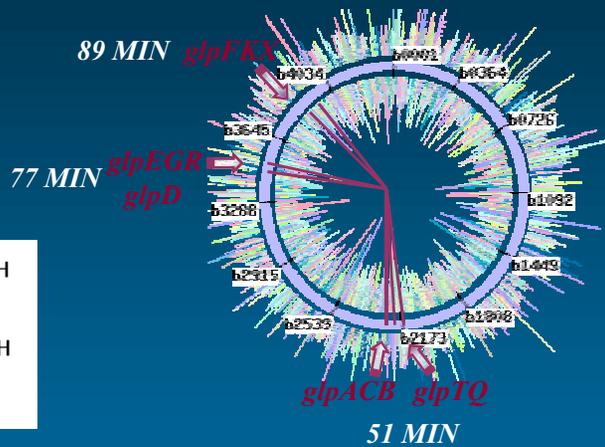
University of California, San Diego  
Department of Bioengineering

Genetic Circuits Research Group  
<http://gcrp.ucsd.edu>

# Glycerol metabolism in E. coli



# Glycerol metabolism related genes in E. coli K12 MG1655

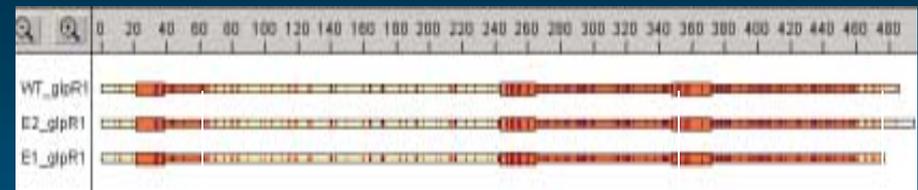


Glycerol: A simple molecule — complex genetic machinery

University of California, San Diego  
Department of Bioengineering

Genetic Circuits Research Group  
<http://gcrp.ucsd.edu>

# MACAW Alignment of glpR-1 wild-type, E1 and E2



60 GUC GUG  
VAL20VAL

350 GUC GGC  
VAL119GLY

Wobble base pair  
No amino acid sequence change  
Found in both strains

535 GGG UGG  
GLY179TRP

University of California, San Diego  
Department of Bioengineering

Genetic Circuits Research Group  
<http://gcrp.ucsd.edu>

## Summary

- ∕ A number of phenotypes with optimality properties have been determined in wt *E. coli* K-12
- ∕ *E. coli* can find new optimal solutions
  - algorithm unknown
  - ~500 to 700 generations, a biphasic process
- ∕ Many biologically equivalent solutions exist
- ∕ *E. coli* ‘forgets’ and appears to have ‘limited memory’
- ∕ There seem to be conflicting optima
- ∕ Knock-out can be evolved towards new optimal functions *a priori* predicted by *in silico* model

## Support

NSF BES-0120363  
NSF MCB-9873384  
NIH R01 GM57089

NSF BES-9814092  
NIH R01 GM62791

Papers can be accessed from..

<http://systemsbiology.ucsd.edu>